

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Hipótesis	3
1.3. Objetivos	3
1.4. Estructura de la tesis	4
2. Revisión de la literatura	5
3. Marco teórico	9
3.1. Aprendizaje reforzado	9
3.1.1. Algoritmo TD3	13
3.2. Aprendizaje reforzado offline	15
3.2.1. Algoritmo CQL	17
3.2.2. Algoritmo IQL	19
3.2.3. Algoritmo AWAC	20
3.2.4. Algoritmo CRR	21
4. Metodología	25
4.1. Simulación	25
4.1.1. Especificación del LHD	26
4.1.2. Especificación del mundo	27
4.1.3. Especificación del controlador	28
4.2. Modelo RL	29
4.2.1. PO-MDP	30
4.2.2. Observaciones	30
4.2.3. Acciones	31
4.2.4. Función de recompensa	31
4.2.5. Estados terminales	33
4.2.6. Exploración y explotación	34
4.3. Planificación y selección de objetivos	34
4.4. Evaluación	37
4.5. Extracción y estudio de datasets	38
5. Resultados	39
5.1. Deep Reinforcement Learning	39
5.2. Generación de datasets	41

5.3. Offline Reinforcement Learning	43
6. Discusión	51
7. Conclusión	57
Bibliografía	57
Anexos	62
Anexo A. Metodología de aceleración en la extracción de datos mediante teleoperación	62
Anexo B. Extracción de dataset de teleoperación en simulación	64